ECON 582 Replication of “Experimental Estimates of Education Production Function” by Allan Krueger

Nikhil Narayane, Vahid Disfani

November 2021

1 Introduction

We are interested in replication of Krueger’s study Krueger, 1999, “Experimental Estimates of Education Production Functions,” published in the *Quarterly Journal of Economics*. The outcome of interest in this study is student achievement and how it is impacted by school input. Krueger hypothesizes that school resources or expenditure affects the student performance. Such a study was really important at the time since (i) large literature on the impact of school resources on students performance were generally ambiguous, conflicting, and weak, and (ii) revealing a clear causal relationship between class size and student performance would have significantly affected the resource allocation in schools and school districts across USA.

To draw strong conclusions, it was very important at the time to provide clear results via a longitudinal studies with large-scale randomized experiments. Thus, Krueger dug into the data of Project STAR as the single largest study conducted in the field of education across USA. In his study, he let his outcome of interest be measured by the students’ performance in a battery of standardized tests given at the end of each school year. To measure the school expenditure, Krueger looked at the class size at 3 different levels: small classes (13-17 students per teacher), regular-size classes (22-25 students), and regular/aide classes (22-25 students) with a full-time teacher’s aide.

There was a significant challenge that Krueger dealt with at the time of conducting this study. In a hypothetical case, let us assume that this effort was not a randomized experiment, but that Krueger were able to measure students’ achievements and school expenditure for a large i.i.d. sample, as well as other control variables. Such assumption would allow for estimation of a least squares regression. In order to interpret the calculated estimators as the causal effect of class size on students’ test results, it would be necessary to assume that conditional on other controls, class size is statistically independent of unobserved variables affecting the students’ test results. The reality was however different from what we just assumed here: in his article, Krueger listed several potential omitted variables that could violate the conditional independence assumption in a hypothetical non-experimental study, and in the non-experimental literature on this particular research question. One example is inconsistency of one's classmate which may influence the achievement. The other example is the non-random transition between classes with different sizes due to behavioral problems and parental complaints. Such parameters *i.e.* classmate inconsistency, behavioral problems and parental interference is generally hard to measure and their impact on students' performance and school expenditure are difficult to quantify.

Krueger also listed another limitation of the experiment which was unavailability of the baseline test score information on the students. Therefore, no one could examine whether the treatment and control groups "looked similar" on this measure before the experiment began. To address these concerns, Krueger argues that a successful random assignment of students between class types will overcome the impact of such omitted variables. For example, if students are successfully randomly assigned between classes, one would expect those assigned to small- and regular-size classes to look similar along other measurable dimensions at base line.

Due to the significance of this study and to obtain a great understanding of instrumental variable (IV) estimation through a hands-on coding effort, we have worked on replicating of Krueger, 1999 using the very same data that Krueger used for hi study in 1999 as explained in the following sections.

2 Data and Replication Steps

Our source data comes from longitudinal study conducted by a team of researchers at Tennessee State University, Memphis State University, the University of Tennessee and Vanderbilt University. This data can be obtained from Achilles et al., 2008. This project which was named Student Teacher Achievement Ratio(STAR) was funded by Tennessee Legislature and was conducted for four years over students entering in kindergarten and exiting in grade 3. The data was collected till 9^{th} grade but–similar to the main reference–this study deals with data till grade 3 which involves 11,600 students over four years. Original data deals with lot of independent variables such as years of experience of a teacher, scaled score of students in math, reading, and vocabulary. Similar to Krueger's article, the outcome of interest for our study (Y) is student performance in the above mentioned tests, which is defined by averaging the individual percentile scores. Through Project STAR, students were tested on two tests such as SAT (Stanford achievement test) and Tennessee BSF (Basic skill first). Since the score has no unit, the researcher converted the test scores into percentiles rank. These percentile ranks were calculated for each class size and for each category of the test.

Krueger showed in his article that the outcome is mainly affected some explanatory variables (X) such as size of the class and teacher aid treatment. Other variables affecting the outcome are free lunch, teacher experience, educational qualifications of the teacher, gender of the teacher. All these variables provide a rich set of covariates to make the dependent variable and other omitted variables independent of each other. The type of these variables are different: some are categorical (*e.g.,* class size and teachers' highest degree); some are binary (*e.g.,* students' and teachers' gender); and some are continuous (*e.g.,* teachers' experience).

To perform the analyses and replicate the selected tables and figure from the paper as instructed, we imported the data into Stata. As always, the first step was to clean the data and make it usable to run the statistical models. Then, we created outcome variable by taking average percentile of students' SAT scores in English, math, and vocabulary. Moreover, we defined number of dummy variables needed in the following regression studies such as the ones indicating class types, students' background, and teachers' characteristics, class type and class size, etc. Finally, performing the regressions as described in the reference led us to the final results we are reporting in this replication report.

3 Design and Implementation

A limitation of this experiment as explained by the author is that baseline test score information on the students is not available. This information is important because to study the causal effect of class size and teacher aid on student performance, we should know how the students performed before they were enrolled in the STAR program. Since we don't know the initial level of difference between different class size students, we can reduce the bias by randomly assigning students to different class sizes. This random assignment will make sure that there is no significant endogeneity in the experiment. Table 1, which exactly matches Table I of Krueger, 1999, reports a comparison of mean characteristics of treatments (class sizes and joint p-value as column variables) and controls (row variables) for all students categorized based on their entry year to the program. The joint p-values are important to study since they test the hypothesis that different class sizes are similar in behaviour. Since the joint p-values are not significant the null hypothesis of different classes behaving similarly because of random assignment of students could not be rejected, and the mean of different groups are statistically equal.

As some of these differences are statistically significant, Table 2 (matching Table II of Krueger, 1999 exactly) is studied to study the importance of controlling for school effects. Since random assignment was only valid within schools, these p-values do not incorporate school fixed effects. Therefore, controlling the school fixed effects while comparing the means of different groups will make our analysis more credible.

Although the results in Table 2 show that there are small differences in the fraction of students on free lunch, the racial mix, and the average age of students in classes of different size, none of the three background variables displays a statistically significant association with class-type assignment at the 10 percent level. It suggests that random assignment produced relatively similar groups in each class size, on average. Within schools, there is no apparent evidence that initial assignment to class types was correlated with student characteristics.

We have defined class sizes with a specific cutoff for each class size, but the data shows that the number of students for a particular class was not fixed around the year. Such irregularity in the number of students were partly due to students moving to different cities, failing a grade, or skipping a grade. This might cause some overlap between different class sizes, possibly leading to some cases where small class size are greater than regular class sizes. To study this, Table III of Krueger, 1999 is replicated in Table 3 which reports the frequency distribution of class size for first graders, by assignment to small, regular, or regular/aide classes. Even though the class sizes may have differed in different schools, the number of students in each class type on average were consistent with the initial design.

Tables 1-3 which show the behaviour of different class sizes and the frequency distribution of class size play very important roles in providing a solid foundation for causal effect of class size on student performance and removes any doubts regarding the credibility of this study.

4 Performance vs. Class Size in a Glance

Figure 1 displays the kernel density of the average test score distributions for students in small and regular classes at each grade level, where the solid line and dashed lines correspond to the performance of the students in small and regular-size classes, respectively. Note that the regular-size classes here include the ones with and without aides. For all grade levels, the results suggest that the students in small classes achieved higher test scores on average compared to the other students. Also, the author points out in Krueger, 1999 that presence of aides in regular-size classes did not significantly affected the students' performance. It is also evident that the highest value achieved by small class size students keeps decreasing as we move from kindergarten to grade 3. This shows that students who entered the program late performed less efficiently than the ones who entered first.

5 Statistical Models

5.1 Ordinary Least Square (OLS) Model

Let us model the effect of school resources on student achievement suing the general model

$$Y_{ij} = aS_{ij} = bF_{ij} + \epsilon_{ij} \tag{1}$$

where the Y_{ij}, S_{ij}, and F_{ij} denote the achievement level of student i in school j, a vector of school j characteristics, and a vector of student i's family background, respectively. We let ϵ_{ij} represent the stochastic error component. The vectors

S_{ij} and F_{ij}–which include the entire history of student's family background and school resources–as well as the student's unobserved inherent ability may impact the student's achievement. Due to lack of such data in practice, these variable are omitted and appear in error term. Th challenge is that if these omitted variables are correlated the included variables, the estimated parameters will be biased. To address this concern, we let the school characteristics be determined by random assignment. With such a random assignment, an unbiased estimate of the effect of class size on achievement is achieved by a simple comparison of mean achievement between students in small and large classes.

For students in each grade level, let us estimate the regression equation

$$Y_{ics} = \beta_0 + \beta_1 SMALL_{cs} + \beta_2 REG/A_{cs} + \beta_3 X_{ics} + \alpha_s + \epsilon_{ics} \tag{2}$$

for student i in class c at school s, where Y denote the average percentile score on the SAT test of the student. The variables $SMALL$, and REG/A are dummy variables indicating whether the student was assigned to a small class or a regular-size class in a specific year, respectively. Also, X contains the vectors of observed student and teacher covariates (e.g., gender). Using a separate dummy variable α_s, we absorb the school effect since the independence of class size and other variables are only valid within schools.

5.2 Two-Stage Least Squares (2SLS) Models

Due to student mobility and enrollment difference across schools, students assigned to small or regular-size classes likely experienced varying number of students in their classes. Krueger, thus, proposes a 2SLS model taking the actual class size into account. This model is formulated as

$$CS_{ics} = \pi_0 + \pi_1 S_{i0s} + \pi_2 R_{i0s} + \pi_3 X_{ics} + \sigma_s + \tau_{ics} \tag{3}$$

$$Y_{ics} = \beta_0 + \beta_1 CS_{ics} + \beta_2 X_{ics} + \alpha_s + \epsilon_{ics} \tag{4}$$

for student i in class c at school s, where CS is the actual number of students in the class, and S and R denote dummy variables indicating initial assignment of students to small and regular-size classes, respectively. The other variables are as defined before. It is obvious that the difference in modeling the OLS (with just dummy variables for small and regular-size classes) and 2SLS (taking the actual size of classes into consideration) leads to different statistical results obtained via these two methods.

In this setup, variation in actual size in test score in (5.2) only includes variations due to the initial assignment to a small or regular-size class in (3). The random assignment of the initial class size then assures that this excluded instrumental variable (actual class size) is uncorrelated with ϵ_{ics}, which is required by 2SLS to be consistent.

6 OLS and 2SLS Estimation Results

To replicate Table V of Krueger, 1999 (summarized in Tables 4-7 in this report), the equation is first estimated by ordinary least squares (OLS). With this structure, the error term ϵ_{ics} will consist of two components: a class-specific random component that is common to all members of the same class (addressing the effects of unobserved teacher characteristics or other students' common influence) and an idiosyncratic error term.

To address the possible impacts of inconsistency of one's classmate or nonrandom transition between classes with different sizes due to behavioral problems and parental complaints, this study also includes another set of dummy variables indicating the students' initial class assignments are defined, and we let them replace the dummy variables of actual class assignment in (2).

The regression results are shown in Tables 4-7 for Kindergarten, first-grade, second-grade, and third-grade students, respectively. These tables consist of 8 columns: columns 1-4 which use actual class assignment and columns 5-8 which use initial class assignment. The school dummies are are omitted in columns 1 and 5 only. These columns show significant gaps in average performance between the students in small classes and those in regular-size classes. These gaps in are 5, 8-9, 5-6, and 5-6 percentile points for kindergarten to third-grade classes respectively. Comparison of the results in columns 1 and 5 shows that nonrandom transition between classes with different sizes did not influence the impact on the impacts of class size on students' performance. Except for first-graders, including school dummy variables in columns 2 and 6 resulted in slight increase in the effect of studying in small classes.

As shown in columns 3 and 7, Students' characteristics have significant explanatory power. There are gaps of 8, 7, 7, and 6 percentile points between White and Asian students' performance and black students' performance. Girls score 3-4 points higher than boys in each group. Students on free lunch score 13 percentile points lower than the others, which likely signals the fact that the students on free lunch come from families with less financial resources. The teacher characteristics, on the other hand, add considerably weak explanatory power: teacher education and gender show systematic effects, and their experience has a small positive effect.

In summary, students in small classes perform better than the others by about 5 percentile points. Adding full-time aides to the regular classes improves the performance slightly. Also, the slight difference between the results reported in columns 1-4 (where the actual assignment is used) with those in columns 5-8 (where the initial assignments are used) suggests that the nonrandom movement of students between classes with difference sizes posed little limitation to the experiment.

Table 8 (replicating Table VIII of Krueger, 1999) presents OLS and 2SLS estimates of (5.2). The 2SLS estimates are slightly larger in absolute value than OLS estimates for Kindergarten, first-grade, and second-grade classes. This difference is significant for third-grade classes. According to the 2SLS estimates, if the size of a class decreases by 10 students the average percentile ranking of students would increase by 8-10 points, depending on the grade. No obvious trend over grade levels in the effect of class size is noticed. It should be noted that there is some discrepancy between the results achieved in Table 8, especially for the kindergarten and first-grade classes, which is likely due to some replication errors and/or difference in data sets.

Q2.G. The report so far has been about following in Krueger's (1999) footsteps. Conclude by describing what you think could be done to improve or expand the analysis without compromising its "internal validity" (meaning, the strength of its causal inferences for this sample of students). Be sure to check the related literature on the STAR program and propose a new improvement or extension.

7 Conclusions and Potential Improvements

This report includes replication of replicating of some selected results presented in Krueger, 1999. A summary of the data used, the replication steps, the experiment design, and statistical models are provided. The obtained results through replication match the original results very well and are explained and discussed in details in this report. From these results, it could be concluded that the size of class had obvious effect on student performance while full-time aide did not. These conclusions remain intact in the modified OLS and the 2SLS estimation processes although the estimates in these two methods are slightly affected by incorporating the initial class size assignments or actual class sizes (instead of the actual class types in the original OLS method).

The author's obtained results match the original results very well and suggests that there is significant causal effect of class size on the performance of student. His study suggests that even though the STAR experiment was not carried flawlessly, it could still be studied because of controlling for covariates. To improve the validity of the study, one thing we can do is to analyse Table III. Table III shows that the three different class size are not distributed uniformly in all the schools. Some schools may have higher number of small size class. To incorporate this effect, we can assign different weights to different class size within the broader category. This is based on Solon's reasoning. Thus, by using, Solon's Weighted GLS variance, the validity of the experiment could be reinforced.

Data Management Plan

Our data management plan included a dedicated directory to manage all the codes, results, reports, and documentations in specific folders. All the data processes were automated. one single execution file **main.do** manages two other **.do** files in order to perform the replication steps. These codes include lines of codes for different tasks in the process including reading data from files, data cleaning, data analysis, visualization, and report generation. One can just run **main.do** file in order to replicate the entire analysis.

For version control, a version of the codes with date/time tags has been generated. The results are all time-stamped as well. To make sure the work is always backed up, the folder specified to the project has been backed up regularly on Github. It could be accessed at https://github.com/nikhilstat/Replication-of-Krueger-study. The final documentations are prepared in LATEX/Overleaf

Tables and Figures

Table 1: Replication of Table I of Krueger, 1999: Comparison of Mean Characterisics of Treatments and Controls: Unadjusted Data

A. Students who entered STAR in kindergarten				
Variable	Small	Regular	Regular/Aide	Joint P-value
Free lunch	0.47	0.48	0.50	0.08
White/Asian	0.68	0.67	0.66	0.26
Age in 1985	5.44	5.43	5.43	0.33
Attrition rate	0.49	0.52	0.53	0.02
class size in kindergarten	15.12	22.38	22.77	0.00
Percentile score in kindergarten	54.72	49.92	50.02	0.00
B. Students who entered STAR in first grade				
Free lunch	0.59	0.62	0.61	0.52
White/Asian	0.62	0.56	0.64	0.00
Age in 1985	5.78	5.86	5.88	0.03
Attrition rate	0.53	0.51	0.47	0.07
class size in kindergarten	15.87	22.71	23.46	0.00
Percentile score in kindergarten	49.49	42.90	47.99	0.00
C. Students who entered STAR in second grade				
Free lunch	0.66	0.63	0.66	0.60
White/Asian	0.53	0.54	0.44	0.00
Age in 1985	5.88	5.91	5.94	0.40
Attrition rate	0.37	0.34	0.35	0.58
class size in kindergarten	15.50	23.71	23.59	0.00
Percentile score in kindergarten	46.65	45.47	41.77	0.01
D. Students who entered STAR in third grade				
Free lunch	0.60	0.64	0.69	0.04
White/Asian	0.66	0.57	0.55	0.00
Age in 1985	5.95	5.93	5.99	0.49
Attrition rate	NA	NA	NA	NA
class size in kindergarten	15.97	24.05	24.43	0.00
Percentile score in kindergarten	47.88	44.53	41.60	0.00

Table 2: Replication of Table II of Krueger, 1999: P-Values for Tests of Whitin-School Differences among Small, regular, and Regular/Aide Classes

Variable	Grade entered STAR program			
	K	1	2	3
Free lunch	0.45	0.29	0.58	0.18
White/Asian	0.66	0.28	0.18	0.26
Age	0.43	0.12	0.43	0.47
Attrition rate	0.01	0.37	0.84	NA
Actual class size	0.00	0.00	0.00	0.00
Percentile score	0.00	0.00	0.49	0.00

Table 3: Replication of Table III of Krueger, 1999: Distribution of Children across Actual Class Sizes by Random Assignment Group in First Grade

Actual class size	Assignment group in first grade			
in first grade	Small	Regular	Aide	Total
12	24	0	0	24
13	182	0	0	182
14	252	0	0	252
15	465	0	0	465
16	256	16	0	272
17	561	17	0	578
18	108	36	0	144
19	57	76	57	190
20	20	200	120	340
21	0	378	378	756
22	0	594	330	924
23	0	437	460	897
24	0	384	264	648
25	0	175	225	400
26	0	130	234	364
27	0	54	108	162
28	0	28	56	84
29	0	29	58	87
30	0	30	30	60
Average class size	15.7	22.7	23.4	

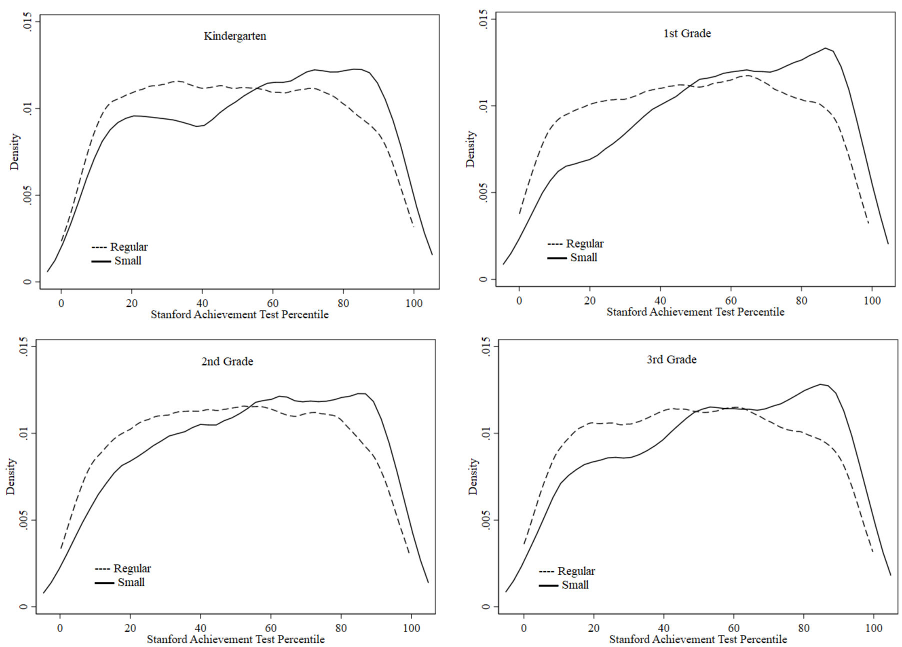

Figure 1: Replication of Figure I of Krueger, 1999: Distribution of Test Percentile Scores by Class Size and Grade

Table 4: Replication of Table V Part A of Krueger, 1999: OLS and Reduced-Form Estimates of Effects of Class-Size Assignment on Average Percentile of Standard Achievement Test or Kindergarten Students

Variable	OLS: actual class size (1)-(4)				Reduced form: initial class size (5)-(8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Kindergarten				
Small class	4.784	5.346	5.393	5.398	4.784	5.346	5.393	5.398
	(1.535)***	(1.410)***	(1.350)***	(1.328)***	(1.535)***	(1.410)***	(1.350)***	(1.328)***
Regular/aide class	0.049	0.202	0.441	0.307	0.049	0.202	0.441	0.307
	(1.272)	(1.195)	(1.172)	(1.148)	(1.272)	(1.195)	(1.172)	(1.148)
White/Asian			8.350	8.464			8.350	8.464
			(1.586)***	(1.604)***			(1.586)***	(1.604)***
Girl			4.393	4.369			4.393	4.369
			(0.568)***	(0.569)***			(0.568)***	(0.569)***
Free Lunch			-13.122	-13.039			-13.122	-13.039
			(0.926)***	(0.934)***			(0.926)***	(0.934)***
White Teacher				-0.675				-0.675
				(2.399)				(2.399)
Teacher Experience				0.263				0.263
				(0.124)**				(0.124)**
Master's Degree				-0.428				-0.428
				(1.074)				(1.074)
R^2	.0066911	.2445137	.3046342	.3079906	.0066911	.2445137	.3046342	.3079906

Table 5: Replication of Table V Part B of Krueger, 1999: OLS and Reduced-Form Estimates of Effects of Class-Size Assignment on Average Percentile of Standard Achievement Test or First Grade Students

Variable	OLS: actual class size (1)-(4)				Reduced form: initial class size (5)-(8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	B. First Grade							
Small class	8.820	8.523	7.897	7.547	7.648	7.111	6.760	6.536
	(1.425)***	(1.367)***	(1.336)***	(1.358)***	(1.260)***	(1.269)***	(1.253)***	(1.252)***
Regular/aide class	3.484	2.246	2.188	1.745	1.895	1.600	1.614	1.483
	(1.351)**	(1.105)**	(1.121)*	(1.156)	(0.910)**	(0.738)**	(0.719)**	(0.722)**
White/Asian			7.002	6.984			6.885	6.857
			(1.443)***	(1.471)***			(1.459)***	(1.488)***
Girl			3.804	3.827			3.769	3.796
			(0.634)***	(0.634)***			(0.623)***	(0.626)***
Free Lunch			-13.427	-13.544			-13.589	-13.709
			(1.201)***	(1.204)***			(1.211)***	(1.213)***
White Teacher				-3.961				-4.047
				(2.510)				(2.535)
Teacher Experience				0.057				0.067
				(0.073)				(0.073)
Master's Degree				0.430				0.579
				(1.248)				(1.269)
School Fixed Effect	No	Yes	Yes	Yes	No	Yes	Yes	Yes
R^2	.0174835	.2351912	.2973702	.2996113	.0129808	.2299106	.2933078	.2959857

Table 6: Replication of Table V Part C of Krueger, 1999: OLS and Reduced-Form Estimates of Effects of Class-Size Assignment on Average Percentile of Standard Achievement Test or Second Grade Students

Variable	OLS: actual class size (1)-(4)				Reduced form: initial class size (5)-(8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	C. Second Grade							
Small class	5.724	6.219	5.789	5.791	5.402	5.684	5.349	5.288
	(1.578)***	(1.422)***	(1.412)***	(1.419)***	(1.366)***	(1.258)***	(1.249)***	(1.252)***
Regular/aide class	1.084	1.394	1.539	1.672	-0.076	1.089	1.238	1.304
	(1.521)	(1.324)	(1.283)	(1.283)	(1.083)	(0.943)	(0.901)	(0.861)
White/Asian			7.055	7.085			6.992	7.025
			(1.556)***	(1.551)***			(1.554)***	(1.551)***
Girl			3.220	3.208			3.221	3.207
			(0.658)***	(0.661)***			(0.654)***	(0.656)***
Free Lunch			-13.631	-13.566			-13.760	-13.693
			(0.766)***	(0.763)***			(0.778)***	(0.773)***
White Teacher				0.487				0.518
				(2.350)				(2.330)
Teacher Experience				0.104				0.108
				(0.080)				(0.080)
Master's Degree				-1.138				-1.243
				(1.265)				(1.276)
School Fixed Effect	No	Yes	Yes	Yes	No	Yes	Yes	Yes
R^2	.0082007	.210267	.2764134	.2781918	.0080247	.2085843	.275147	.2768114

Table 7: Replication of Table V Part D of Krueger, 1999: OLS and Reduced-Form Estimates of Effects of Class-Size Assignment on Average Percentile of Standard Achievement Test or Third Grade Students

Variable	OLS: actual class size (1)-(4)				Reduced form: initial class size (5)-(8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				D. Third Grade				
Small class	5.660	5.788	5.118	5.115	5.908	5.629	5.346	5.389
	(1.481)***	(1.286)***	(1.261)***	(1.286)***	(1.214)***	(1.145)***	(1.077)***	(1.133)***
Regular/aide class	-0.365	-0.094	-0.234	-0.503	-0.353	0.161	0.113	0.058
	(1.447)	(1.263)	(1.307)	(1.273)	(0.991)	(0.827)	(0.797)	(0.772)
White/Asian			6.108	6.047			5.966	5.906
			(1.551)***	(1.572)***			(1.537)***	(1.563)***
Girl			4.183	4.143			4.199	4.159
			(0.767)***	(0.764)***			(0.766)***	(0.761)***
Free Lunch			-13.197	-13.068			-13.384	-13.263
			(0.929)***	(0.916)***			(0.941)***	(0.931)***
White Teacher				0.372				-0.050
				(1.463)				(1.458)
Teacher Experience				0.058				0.045
				(0.069)				(0.069)
Master's Degree				0.898				0.705
				(1.285)				(1.292)
School Fixed Effect	No	Yes	Yes	Yes	No	Yes	Yes	Yes
R^2	.0103511	.1699918	.2240221	.2235421	.0099623	.1683241	.2234912	.2228216

Table 8: Replication of Table VII of Krueger, 1999: Exploration of Effect of Attrition Dependent Variable: Average Percentile Score on SAT

Grade	OLS	2SLS	Sample Size
	(1)	(2)	(3)
K	-.76	-.81	2,927
	(.14)	(.15)	
1	-.95	-.97	4,125
	(.13)	(.15)	
2	-.61	-.78	4,963
	(.12)	(.14)	
3	-.61	-.82	6,124
	(.13)	(.15)	

References

Achilles, C., Bain, H. P., Bellott, F., Boyd-Zaharias, J., Finn, J., Folger, J., Johnston, J., & Word, E. (2008). *Tennessee's Student Teacher Achievement Ratio (STAR) project*. https://doi.org/10.7910/DVN/SIWH9F

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, *114*(2), 497–532.